CHAPTER 4
DATA SUMMARIZATION AND CLASSIFICATION

4-1. Data transformation

In order to eliminate inaccuracies in the database, the transition from raw data to completed, analyzed data, needed to proceed smoothly. This meant that the data collection effort and the data summarization needed to compliment one another.

   a. Data summarization was very important in that it was the process that transformed the raw data collected at a facility into a standardized format that could be used to develop the reliability and maintainability metrics associated with the components being tracked. If there were difficulties transcribing the raw data during the summarization process then inaccuracies would certainly develop.

   b. Realizing that it was essential to provide an accurate database, Alion personnel needed to identify and implement procedures that would enhance the data summarization process prior to collecting the data. The data collectors would know exactly what data, regarding the components, they would be looking for. The data categories that Alion personnel would focus on were identified as:

   (1) Nameplate information including manufacturer, model number, serial number and size.

   (2) Installation date of component.

   (3) Operation profile of how many hours per day, month, or year the component is required to operate.

   (4) Record of preventative maintenance intervals. If records of completion with dates are not available, how stringent is the site at carrying out these tasks to the prescribed interval.

   (5) Dates of all other events with descriptions that provide the ability to identify the event as a failure or preventative maintenance action.

   (6) If available, the time needed to perform the maintenance event whether preventative or corrective maintenance due to failure.

   c. To ensure that they gathered the most detail on the maintenance events the team would bring a photo copier with them to copy all of the hard copy records from the site. This included maintenance performed by the facilities personnel and also any outside contractors. Fortunately, outside contractor records are typically kept for several years at a facility for warranty issues.

   d. Other standards developed were definitions for components and defining the difference between a preventative maintenance action and a failure. These definitions can be seen in the Glossary and their calculations can be seen in table 5-1.

   e. Once the data was collected, then the summary process began. The first decision that would be made upon reviewing the data was to determine which items had enough complete data to conduct an accurate summarization. As with every data collection program, there are varying degrees of completeness in the data gathered. Due to time restraints during the visit, it is

impossible to review all of the data to determine its completeness. It was determined during the summary process that a total of twenty-eight (28) of the one hundred sixty-two (162) facilities, which were visited, had data which was subsequently rejected due to a variety of deficiencies.

4-2. Data issues

During any data collection effort the goal is always to gather specific information on the component that will be needed for a complete analysis. However, it is often difficult to determine during the collection effort if there are "holes" in the information. It is not uncommon to find out during the summarization process that a complete, thorough time line of information on a component is not documented in the data. Data quality would certainly influence the confidence level that the end user, the facility engineer, would have in the data.

   a. Variations on completeness of data were commonly discovered at this phase. Some data sources had complete records and could give statistics on operational characteristics on every piece of equipment from installation date to that current moment of time. Other sources only recorded a date and a very brief description of an event making it difficult to classify the event as a preventative maintenance task or a corrective action due to a failure.

   b. Analysts were required to sort through numerous maintenance records on components one sheet at a time. Every event would be entered into a spread sheet and a determination based on the data would be made to classify an event as a preventative maintenance task or a failure. Metrics such as time to repair or total time the equipment was down (unavailable) would also be documented if the information was available.

   c. Most of the components that had the greatest detail on maintenance events were the components that served to be most vital to the facilities operations, such as cooling towers and boilers. Smaller, less expensive components, such as valves and filters were typically not tracked by the facility either due to the fact that they did not fail often or were not deemed vital to facility operations. Therefore, the amount of data points on these smaller types of components was limited.

   d. Other problems included incomplete or non-current versions of the equipment's blueprints. Technicians manually developed parts lists, recording data from nameplates and relying on facility engineers for component descriptions. Data on components that did not fail frequently and were not recorded by facility personnel was sometimes developed based on verbal information from experienced technicians.

   e. An example of a situation like this would be if the technician had been employed at the facility for eight years and could remember that he only had to perform maintenance on a particular type of valve three times. The analyst could then make a fairly accurate assessment of the component based on the technician's judgment. In this hypothetical example, however, the analyst would have to determine if this technician would be the only individual that would have made the repair or would have knowledge of the repair.

   f. Due to all of these varying degrees of data completeness, it became apparent that analyzing personnel needed to categorize the different data quality levels to ensure that the final data collection included fair data representation for each component. To quantify this data completion (or quality) index, the technicians classified the data into four levels; perfect data, not perfect

data, verbal/inspection data, and soft data. The distribution of the data collected at the various levels is shown in figure 4-1 and are defined as follows:
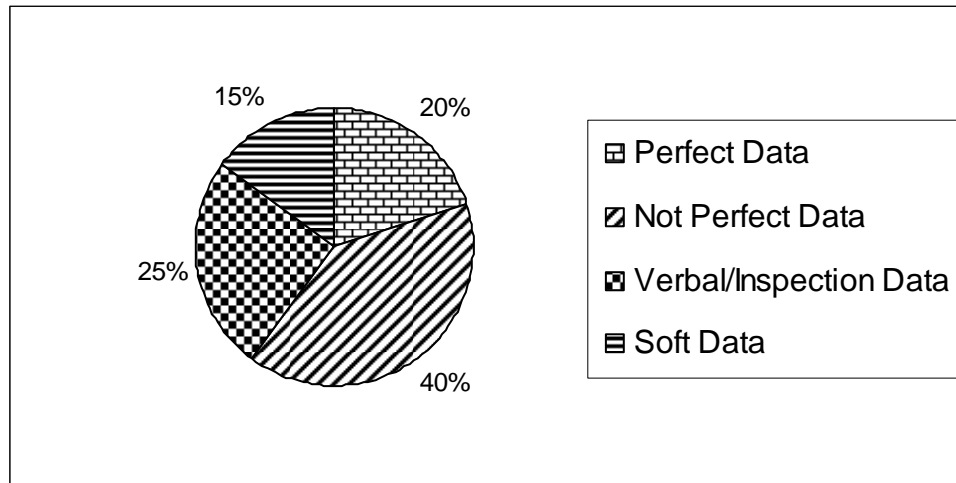

Figure 4-1. Distribution of data quality levels

(1) Perfect Data: Data needed for a valid, complete reliability study, including a parts list, failure history data with time-to-failure statistics, parts description data, operational periods, and ten continuous years of recorded data. No engineering judgment or data extrapolation is required. The PREPIS equipment record database is comprised of 20% of this type of data.

(2) Not Perfect Data: Data with no serious flaws, but the data collection process demanded additional time to ensure useful information was gathered. Examples include parts list determined by inspection, incomplete blueprints or less than ten years of data. The PREPIS equipment record database contains 40% of this type of data.

(3) Verbal/Inspection Data: Data with serious gaps that required additional documentation and verification prior to its inclusion in the database. Items included were typically major items, such as generator sets and boilers. Senior maintenance personnel were interviewed to extract the necessary information to fill the data gaps. These interviews were used as support documentation to recorded data, not as data source information. About 25% of this type of data exists in the PREPIS equipment record database.

(4) Soft Data: Data that relied on the memories of experienced maintenance personnel from the participating facility; it was often extracted from log books containing maintenance personnel entries, filing cabinets with work order forms, and repair records when outside repair support was needed. Engineering judgment was often used to determine numerous performance parameters. This type of data was the most difficult and time consuming to summarize and was only used when no other data sources were available. The PREPIS equipment record database is comprised of 15% of this type of data.

4-3. Maintenance policies and classifications

Maintenance policies and practices have a major influence on reliability and longevity of components. Different facilities, depending on numerous factors, have a variety of approaches to maintenance. Therefore, if data was collected only at facilities with a high degree of

maintenance, then the reliability values would only be accurate for those types of facilities. This information would not be pertinent to facilities with minimal maintenance practices.

a. The major intent of the data collection effort was to minimize the effects of maintenance policies and procedures on the calculated availability values. This was accomplished by collecting data from a variety of locations having various maintenance policies. Technical personnel developed a code to categorize each facility's maintenance policies and procedures into one of three levels:

(1) Code "1": *Above average* maintenance policy. The facility not only followed a scheduled, preventative maintenance policy that was equivalent or similar to the manufacturer's suggested policy, but also went beyond it, such as using redundant units, specialized equipment tests (thermograph, vibration analysis, oil analysis), and complete spare parts kits for equipment.

(2) Code "2": *Average* maintenance policy. Facility used either in-house maintenance crews performing scheduled, preventative maintenance according to the equipment manufacturer's suggested PM schedule or a combination of in-house maintenance crews and outside contractors. In both cases, it was verified that they did follow a fairly rigid schedule.

(3) Code "3": *Below average* maintenance policy. Facility's actual policy was less than average. It may have instituted a scheduled maintenance policy but not followed it or it may have had no maintenance policy. Symptoms such as leaky valves with rags tied around them, dirty air filters, squeaky bearings, loose belts, and general house keeping because of unavailable manpower were typical signs that maintenance at a facility was less than desirable.

b. Each location was then compared to each other and to the average maintenance policy. An overall viewpoint of the different types of maintenance policies that Alion personnel visited can be seen in figure 4-2. This clearly shows that the majority of the facilities that the technicians visited practiced an average level of maintenance; that is, they adhered to the manufacturers recommended maintenance policy. It is also indicates that approximately the same number of facilities that had below average maintenance policies as those facilities that had an above average maintenance policy were data contributors.

Figure 4-2. Comparison of maintenance policies